Global Night Time Lights Urban Extents and Growth Patterns Product[1]

(Alpha Version)

Technical Documentation

Prepared by Mark Roberts, Benjamin Stewart,
Mihir Prakash and Katie McWilliams

GSURR & ITSQS
The World Bank

This draft: 9th November, 2015

[1] This product is part of the "Spatial Patterns of Development in SAR" project (P153150), which is a child activity of the Global Research Program on Spatial Development of Cities (P143985). The development of the product was undertaken by a team led by Mark Roberts and Benjamin Stewart, and also consisting of Katie McWilliams and Mihir Prakash. The product builds on work that had earlier been performed for the SAR Urbanization Flagship (P133593). The team is grateful to Ming Zhang, Peter Ellis, and Paula Restrepo Cadavid for their feedback on an earlier draft of this documentation. They are also grateful to the numerous colleagues and outside collaborators with whom they have had conversations related to the use of night-time lights data which have helped to inform the development of this product. Particular thanks in this respect are due to Klaus Hubacek, Elisa Muzzini, Christopher Small, Luis Quintero, and Naijun Zhou.

1. Introduction

As part of the recently launched South Asia Urbanization Flagship report, *Leveraging Urbanization in South Asia: Managing Spatial Transformation for Prosperity and Livability*,[2] night-time lights (NTL) data collected by satellite were utilized to help delineate the spatial extents of urban areas and analyze spatial patterns of both urban and economic growth. [3] NTL data was also used as an input into a composite index – the Prosperity Index – which measures differences in success across sub-national areas in generating prosperity.[4] The work under the Urbanization Flagship was, in turn, itself inspired by academic research into the use of NTL data both to construct estimates of GDP growth where GDP data is either of poor quality or entirely absent (Henderson, Storeygard and Weil, 2011, 2012; Nordhaus and Chen, 2014) and to monitor spatial patterns of expansion of urban area at a national or regional scale (following, *inter alia*, Zhang and Seto, 2011).[5]

This project builds on the work originally conducted for the SAR Urbanization Flagship, much of which was performed in collaboration with CIESIN[6] and the Department of Geographical Sciences, University of Maryland,[7] to provide a global NTL urban extents and growth patterns product that covers the time-period 1996-2010.[8] For each country, this product takes the form of a basic "data packet" which consists of: ***(i)*** a set of standardized maps which depict the evolution of a country's urban extents and growth rates of NTL intensity at a granular level; and ***(ii)*** a spreadsheet of key output variables relating to, amongst other things, the area and growth of urban extents, as well as their overall "brightness." The product also includes a PowerPoint presentation which is designed to provide an introductory guide to NTL data and the uses to which it has been put in the analysis of both urban and spatial patterns of development. The presentation furthermore outlines the main strengths and limitations of the data, and provides some guidelines on its appropriate use.

The Global Night Time Lights Urban Extents and Growth Patterns product is intended as a base product for use by World Bank teams who wish to perform analysis similar to that reported in the Urbanization Flagship report for countries outside of South Asia.[9] Although a free-standing

---

[2] See http://www.worldbank.org/southasiacities for the full report.

[3] NTL data are collected as part of the United States' Defense Metrological Satellite Program using its Operational Linescan System (DMSP-OLS) sensors. Several different NTL products are available for free download from the website of the National Oceanic and Atmospheric Administration (NOAA): http://ngdc.noaa.gov/eog/dmsp.html. The Global Night Time Lights Urban Extents and Growth Patterns product relies primarily on the global radiance calibrated product.

[4] Chapter 2 of the SAR Urbanization Flagship report describes this index and its method of construction in detail.

[5] Recognition of the ability of NTL data to detect patterns of human activity dates back to Croft (1978).

[6] Center for International Earth Science Information Network, Columbia University – New York.

[7] The work in collaboration with the University of Maryland has subsequently been published as Zhou, Hubacek and Roberts (2015).

[8] 1996 corresponds to the earliest year for which global radiance calibrated NTL data is available.

[9] The project of which this product is an output has already contributed to, amongst others, the Ukraine Urbanization Review (P149719), the fifth edition of the Uganda Economic Update titled "Making Uganda's Cities Fuel the Growth Engines" (P151592), the activity "Urbanization Trends and Spatial Patterns" for Argentin (P154924), and the "Addressing Shrinking Cities in ECA" project (P154478) It has furthermore provided inputs into the Systematic Country Diagnostic (SCD) exercises for, *inter alia*, Bangladesh, Mozambique, Panama, and Sri Lanka. It is anticipated that the project will provide further support into both on-going and planned SCDs for other countries,

product that can be used by itself, it also compliments other global urban and spatial products that are being developed both by the Bank[10] and other institutions. The spreadsheet of key output variables can furthermore be used as an input into further statistical and econometric work on patterns of urban and spatial development. Overall, the product is best suited to providing a broad overview of spatial patterns of urban growth and expansion within a country. Owing to various measurement errors associated with the NTL data,[11] it should not be taken as providing precise quantitative results for any one given urban area.

This technical documentation accompanies the alpha version of the Global NTL Urban Extents and Growth Patterns product, which is open to revisions in the future based on both user feedback and technical advances in methodology. The documentation provides a brief overview of both the product and the methodology which underpins it. It is assumed that the reader of the document has some basic familiarity with NTL data and awareness of its uses without necessarily having had direct prior experience of working with the data itself. For a more comprehensive introduction to NTL data, the reader is referred to CIESIN's excellent, if now slightly dated, "Thematic Guide to Night-time Light Remote Sensing and its Applications." (Doll, 2008).

The remainder of this document consists of four sections. Section 2 describes the contents of the country specific data packets that form the basis of the global product. Section 3 provides a brief explanation of the methodology that underlies the derivation of the country specific NTL thresholds that are used to demarcate urban areas. Finally, Section 4 discusses some of the limitations and future potential extensions and refinements of the alpha product.

---

as well as acting as a more general resource to support the analytical and policy dialogue work of teams across the Bank.

[10] For example, the Standard Urban Definitions global product that is being developed under the Spatial Footprint and Urban Form of Cities project (P149790).

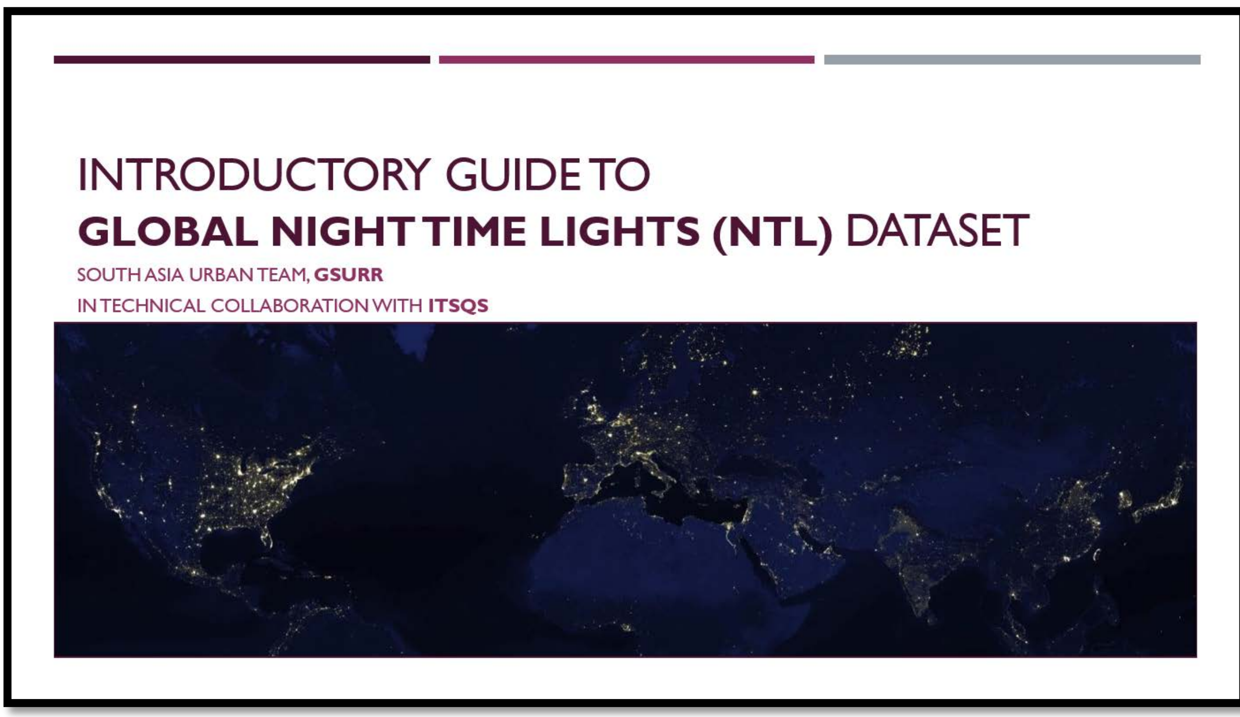
[11] The measurement errors arise from several sources including the relative coarseness (by modern remote sensing standards) of the DMSP-OLS NTL data, the problem of over-glow (otherwise known as "blooming"), and the difficulties associated with performing accurate inter-annual calibration.

2. Contents of Country-Specific Data Packets

The data packet that is available for each country as part of the global urban extents and growth product consists of: ***(i)*** a PowerPoint presentation which provides an introductory guide to global NTL data and its use in urban and spatial analysis; ***(ii)*** a map which shows the evolution of urban extents, as derived from NTL data, between 1996 and 2010; ***(iii)*** a map which shows growth patterns of night-lights at the pixel level between 1996 and 2010; ***(iv)*** a map which shows growth patterns of night-lights within the defined urban extents between 1996 and 2010; and ***(v)*** an Excel spreadsheet which contains output data on key variables relating to the derived urban extents and their growth, and which can be used as input into further statistical and econometric analysis.[12] More detail on each of these five sub-products is provided below.

1.1 PowerPoint: Introductory Guide to Global NTL Data

Figure 1: Screenshot of title slide from "Introductory Guide to Global Night Time Lights (NTL) Dataset" presentation



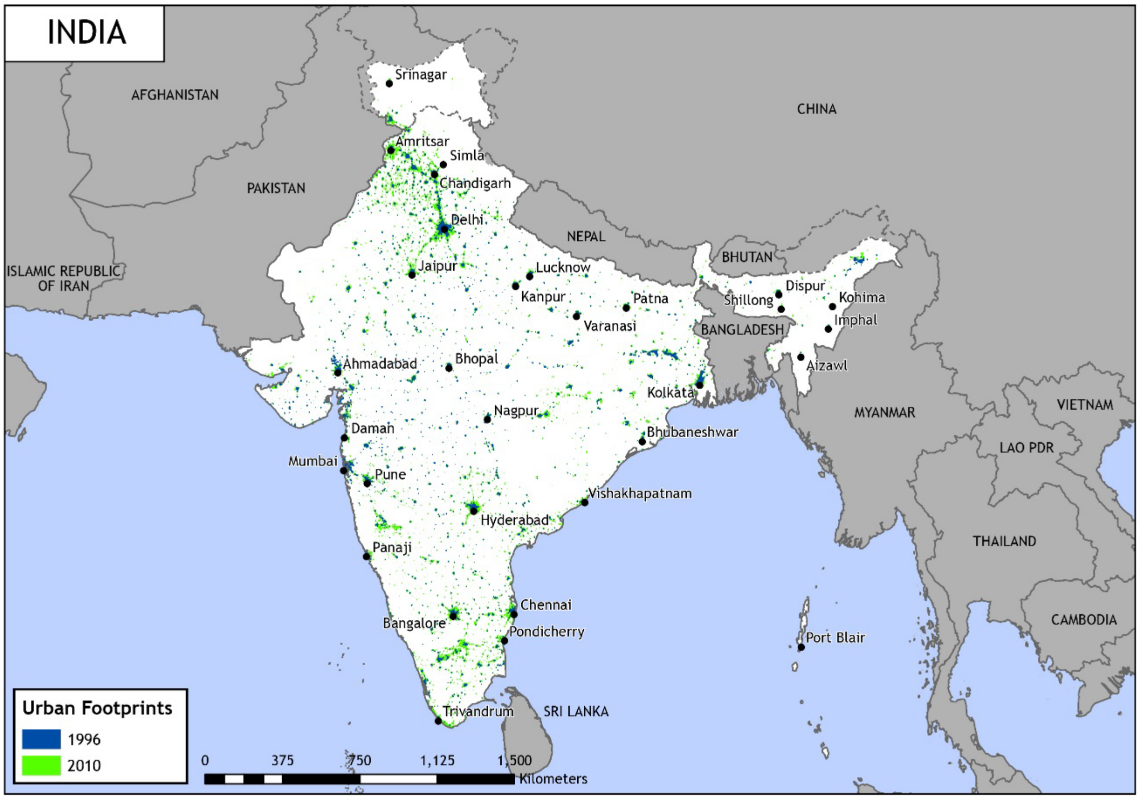
This presentation is designed to introduce users with relatively little or no prior knowledge of NTL data with the available DMSP-OLS night-lights data sets, and to provide basic awareness of how it can be applied to urban and spatial analysis. In this sense, the presentation is a complement to, for example, CIESIN's much more detailed "Thematic Guide to Night-time Light Remote Sensing and its Applications." (Doll, 2008). More specifically, the presentation offers a brief introduction to the source of the data, the different NTL data products that are

[12] Items (ii) – (iv) are all automatically generated outputs of the algorithms that underlie the global urban extents and growth product. The code was all written in Python, and is available upon request.

available for public download from NOAA,[13] and how, drawing on examples from South Asia and beyond, NTL data can be used in urban and spatial analysis. It also provides an overview of the strengths and weaknesses of DMSP-OLS data, and provides some guidelines on its appropriate use in urban and spatial analysis.

1.2 Map: NTL Derived Urban Extents (1996 and 2010)

Figure 2: Output map for NTL derived urban extents - Example of India



The output map depicts urban extents in 1996 and 2010 as derived using the global radiance calibrated NTL data set. This data set contains information on the detected intensity of NTL for grid cells or pixels that have an approximate spatial resolution of 0.86 km^2 at the equator.[14] Intensity of light is measured on a so-called digital number (DN) scale. On this scale, larger DN values corresponding to "brighter" lights. DN values are reported as an annual average over all cloud-free nights within a given year. For both 1996 and 2010, the boundaries of the urban extents depicted in the output map are defined by a country-specific DN threshold using a methodology that is outlined more fully in section 3 of this document. The map provides a visualization of the spatial expansion of detected urban extents between 1996 and 2010. It is

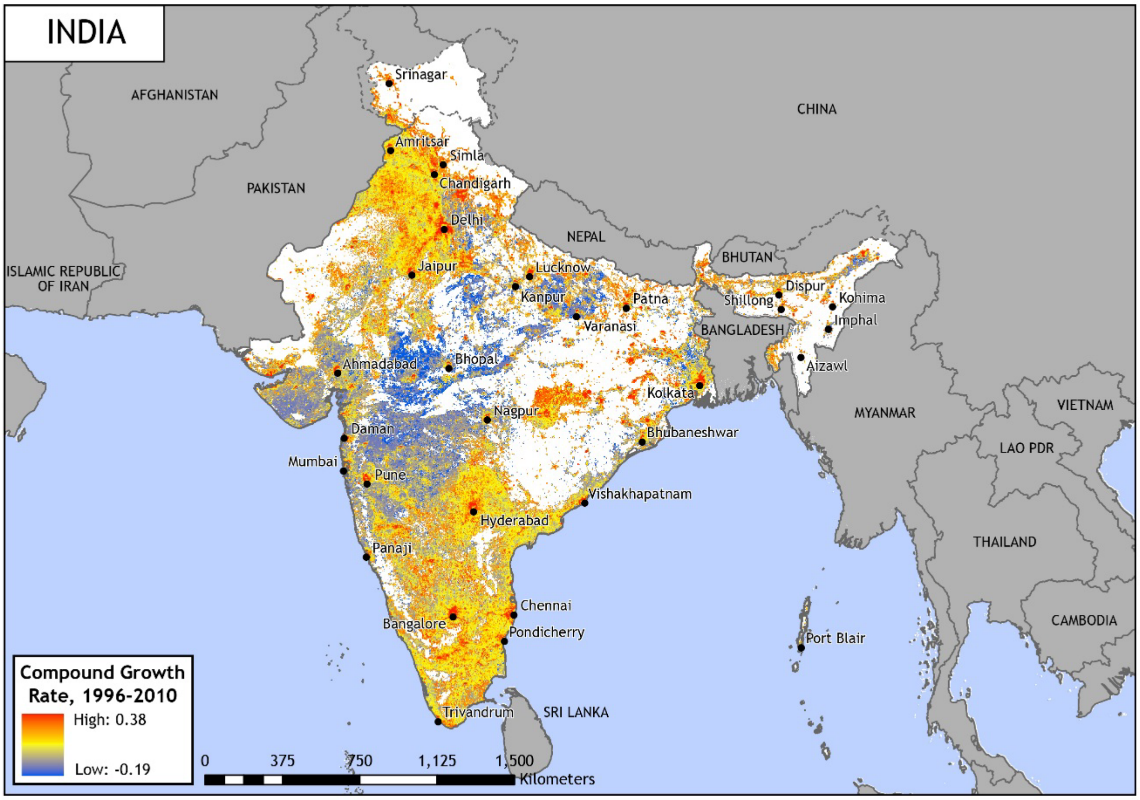
[13] See http://ngdc.noaa.gov/eog/dmsp.html.

[14] Higher-resolution NTL satellite data from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day-Night Band (DNB) sensor mounted on the Suomi National Polar-orbiting Partnership satellite series has recently become available. In particular, monthly composite of this data can be downloaded from the NOAA website at http://ngdc.noaa.gov/eog/viirs.html. These monthly composites, however, only go back as far as April 2012. As such, the data cannot be used to perform long-run analysis of urban and spatial growth patterns.

important to note that these urban extents are unlikely to conform to official administrative boundaries for cities. In broad terms, two types of urban extents can be distinguished – "stand-alone city" urban extents and "multi-city agglomeration" urban extents. As their titles suggest, a "stand-alone city" urban extent is uniquely associated with one city, whereas a "multi-city agglomeration" urban extent covers an area which includes at least two cities (see below for a more detailed discussion of these concepts).

1.3 Map: Compound Annual Growth Rate of NTL Intensity, 1996-2010

Figure 3: Output map for Compound Annual Growth Rate of NTL Intensity, 1996-2010 - Example of India



The output map shows compound annual growth rates of NTL intensity between 1996 and 2010 at the pixel level. These growth rates are calculated using the formula:

$$g_{i,1996-2010} = \left[\left(\frac{DN_{i.2010}}{DN_{i,1996}}\right)^{1/14} - 1 \right] \times 100$$

where $DN_{i,t}$ denotes the DN value for pixel *i* in year *t*.

Although care of interpretation is required, academic literature has presented evidence to show that, at the cross-country level, growth rates of NTL intensity have a statistically significant positive correlation with growth rates of real GDP.[15] This has led to NTL intensity growth rates being used as a proxy indicator of economic growth in situations in which GDP data is either of poor quality or missing entirely (Henderson, Storeygard and Weil, 2011, 2012). This includes at the sub-national level where national statistical offices, particularly in developing countries, frequently do not produce GDP data, especially for lower tier administrative units. This being the case, the output map can be interpreted as providing a broad national-level overview of spatial patterns of economic growth. Although the growth rates depicted in the map should not be interpreted as direct estimates of real GDP growth rates, they can (under appropriate conditions) be interpreted as being positively linearly related to these growth rates.[16] Growth of NTL intensity can also be interpreted as being a reflection of improvements in infrastructure[17] or progress in electrification (see, for example, Min and Kwawa, 2014).

As with the urban extents map, the growth rates are derived using the global radiance calibrated NTL data set. A particular advantage of this product over the more regularly used average visible, stable lights product in this context is that it does not suffer from the so-called "top-coding" problem, whereby variations in the intensity of NTL beyond DN = 63 cannot be observed. As a result, unlike the average visible, stable lights product, the radiance calibrated product allows for growth of bright urban cores to be observed.

1.4 Map: Compound Annual Growth Rate of NTL Intensity within Urban Extents, 1996-2010

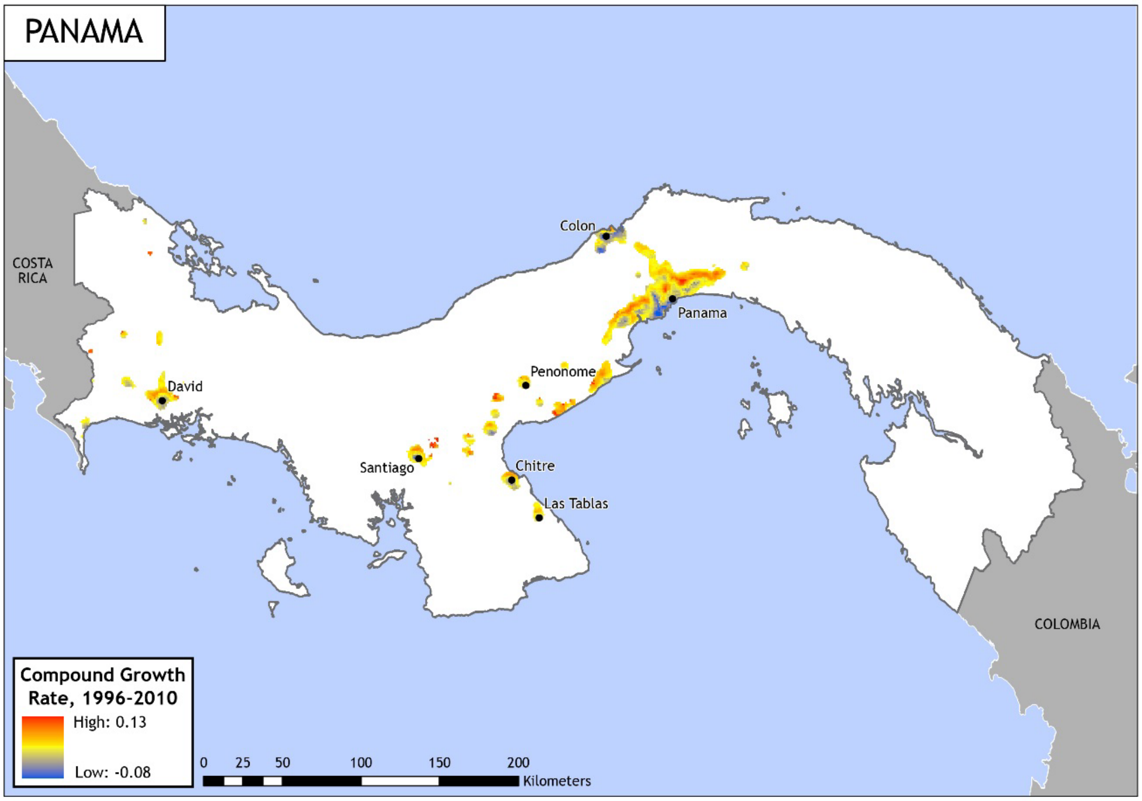
This map is similar to the previous one except that compound annual growth rates of NTL intensity between 1996 and 2010 are only shown within the areas of the 2010 urban extents. Although, just as with the previous map, some care of interpretation is required, it provides potential insights into patterns of development intensification, economic growth or infrastructure improvement within urban areas. In the example shown in Figure 4, rapid growth of NTL intensity can be observed, for instance, in the area around Panama City, while growth at the "core" of the city has, in fact, been negative, thereby indicating a slight "dimming" of lights.

[15] The South Asia Urbanization Flagship report also presents evidence of a statistically significant correlation for a limited sample of Indian districts over the period 1999-2010 (see, in particular, Annex 2A of the report).

[16] The appropriate conditions are that there exists a positive, and stable, underlying linear relationship between real GDP growth and the growth of NTL intensity. Although, in the absence of GDP data, the existence of such a relationship cannot be directly tested (indeed, there would be no need to test it if GDP data were available), it is best practice to test the relationship using GDP data from a higher level of spatial aggregation. The existence of a significant and stable relationship then helps to provide reassurance for the use of NTL intensity growth as a proxy measure of economic growth at more refined spatial levels.

[17] This is how the lights data is being interpreted in, for example, work on African cities under the "Spatial Development of African Cities" (P148736) activity.

Figure 4: Output map for Compound Annual Growth Rate of NTL Intensity within 2010 Urban Extents, 1996-2010 – Example of Panama



1.5 Excel Spreadsheet: Summarized Urban Extents

The spreadsheet contains data on key variables for the derived urban extents for the country, as well as some simple default charts which are automatically generated from the data. More specifically, the spreadsheet consists of the following four worksheets:

- **"Data dictionary":** provides definitions for each of the variables contained in the rest of the spreadsheet.
- **"Extents":** contains information on key variables for each of the identified urban extents within the country.
- **"Cities":** identifies the individual cities that are associated with each urban extent listed in the "Extents" worksheet. Also provides related information on the population of each city (in *circa* 2000) and its location (i.e. latitude and longitude).
- **"Charts":** provides some simple default charts showing patterns of NTL growth across the country's ten most populous urban extents.

The main unit of observation generated by the Global Night Time Lights Urban Extents and Growth Patterns product is intended to be the urban extent. Table 1, therefore, provides an overview of the key variables contained in the "Extents" worksheet.

Table 1: Definition of key variables in the "Extents" worksheet

COLUMN	DEFINITION
EXTENTNAME	Name of urban extent if found
EXTTYPET0	Type of extent in 1996 (Agglomeration, Stand-alone city, -1 is found but without an intersecting city, blank indicates there was no comparable extent)
CTYCNTT0	Number of cities in urban extent in 1996 (only greater than 1 for agglomerations)
EXTTYPET1	Type of extent in 2010 (Agglomeration, Stand-alone city, -1 is found but without an intersecting city, blank indicates there was no comparable extent)
CTYCNTT1	Number of cities in urban extent in 2010 (only greater than 1 for agglomerations)
STATUS	Status of urban extent - Missed (does not intersect any cities) Found (intersects cities in both time periods) Disappear (intersects extent only in 1996) Appear (intersects extent only in 2010)
GAREAKM	Area of urban extent in 2010 in km2
POP	Population tabulated from GRUMP data (~2000)
RC1996_T0	Total brightness of 1996 radiance calibrated NTL for urban extent in T0
RC2010_T1	Total brightness of 2010 radiance calibrated NTL for urban extent in T1
NTLCHANGE	Change of brightness in urban extent (RC2010_T1 - RC1996_T0)
NTLCHGCORR	Change in brightness of urban extent corrected for existing light in expanded area (RC2010_T1 - RC1996_T0 - (RC1996_T1 - RC1996_T0))
INTENSIVE	Change of brightness in urban core (RC2010_T0 - RC1996_T0)
EXTENSIVE	Change of brightness in newly created area of urban extent (RC2010_T1 - RC2010_T0)
EXTENCORR	Change of brightness in newly created area of urban extent corrected for existing light in expanded area (RC2010_T1 - RC2010_T0 - (RC1996_T1 - RC1996_T0))
AREACHG	change in area from 1996 to 2010 in km2
RAWXXXX	Total brightness in raw NTL in year XXXX for T1 extent
RCXXXX	Total brightness in radiance calibrated NTL in year XXXX for T1 extent

Whether or not an urban extent is classified as a "Stand-alone city" or "Agglomeration" under the "EXTTYPET0" (i.e. type of extent in 1996) and "EXTTYPET1" (i.e. type of extent in 2010) variables depends, in part, on the list of cities which is used as input into the algorithms that underpin the creation of the global product. The default input city list for a country is provided by CIESIN's GRUMP urban settlement point layer.[18] An urban extent is, therefore, categorized as a "Stand-alone city" if only a single settlement point from this layer falls within its area.[19] By contrast, an urban extent is an "Agglomeration" if more than one settlement point from the GRUMP layer falls within its area. The "Extents" worksheet also records instances of urban extents that do not intersect with a settlement point from the GRUMP layer (in these cases the variable "EXTTYPET0" or "EXTTYPET1" takes on a value of -1).[20]

[18] GRUMP is an acronym for Global Rural Urban Mapping Project. This project includes a database that maps populated places across the globe: http://sedac.ciesin.columbia.edu/data/set/grump-v1-settlement-points.

[19] To take account of geo-locational errors the underlying Python code defines a small buffer zone around each urban extent. Provided a settlement point falls within an urban extent's buffer zone, it is classified as belonging to that urban extent.

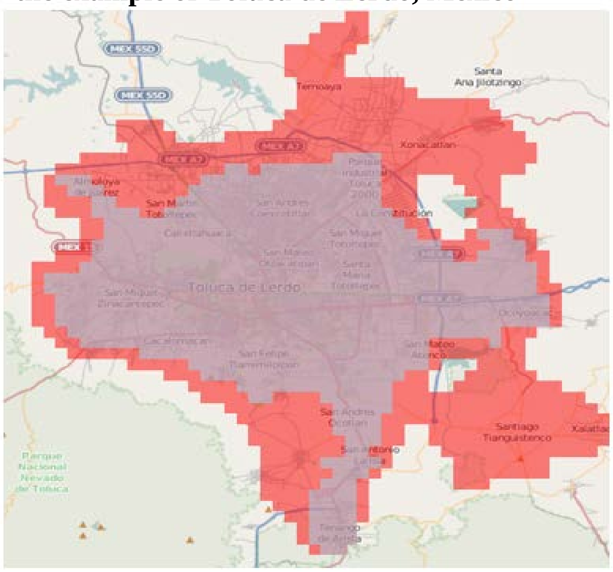
[20] These "unclassified" urban extents may be worthy of further analysis in themselves. Thus, they may, for example, represent interesting instances of urbanization that is occurring "under the radar." Alternatively, the user

Meanwhile, the variables "INTENSIVE", "EXTENSIVE", and "EXTENCORR" provide information on the type of growth being experienced by a stand-alone city or multi-city agglomeration. Box 1 explains in more detail the derivation of these variables and their meanings.

Box 1: Describing patterns of intensive and extensive growth

An urban extent can grow in two basic ways: ***(i)*** intensively; and/or ***(ii)*** extensively. In the Global NTL Urban Extents and Growth Patterns product, intensive growth corresponds to the situation in which the area within the original (1996) urban extent becomes brighter overall during the period 1996-2010. Thus, in the example of Figure B1, the area shaded blue, which shows the 1996 urban extent for Toluca de Lerdo, Mexico, becomes brighter overall. This increase in overall brightness is measured by the variable "INTENSIVE" in the "Extents" worksheet. Meanwhile, extensive growth occurs when the area of the urban extent grows and corresponds to the additional brightness in the newly added urban area (i.e. to the additional brightness in the red shaded area in Figure B1). Two different measures of extensive growth are reported in the "Extents" worksheet. The first, "EXTENSIVE", ignores the pre-existing brightness in the newly added area, whilst the second, "EXTENCORR" takes account of this pre-existing brightness. More detailed technical definitions of the variables "INTENSIVE", "EXTENSIVE", and "EXTENCORR" are provided below.

Figure B1: Intensive and extensive growth – the example of Toluca de Lerdo, Mexico



Let $DN_{e,t}$ = DN value in extent e in year t where $e \in \{t0, t1\}$ and $t \in \{1996, 2010\}$. Hence, for example, $DN_{t0,2010}$ denotes the total intensity of NTL in the $t0$ extent in 2010. More generally, NTL intensity values for different combinations of extent and year are as given in the table below:

Extent / Year	1996	2010
$t0$	$DN_{t0,1996}$	$DN_{t0,2010}$
$t1$	$DN_{t1,1996}$	$DN_{t1,2010}$

Given this notation, it follows:

NTLCHANGE	**= Total change in NTL intensity**	$= DN_{t1,2010} - DN_{t0,1996}$
INTENSIVE	**= Intensive Change**	$= DN_{t0,2010} - DN_{t0,1996}$
EXTENSIVE	**= Extensive Change**	$= DN_{t1,2010} - DN_{t0,2010}$

---

may wish to drop these urban extents from further analysis if she is only interested in those that have associated identified urban settlement points.

On the above definitions, the total change in NTL intensity is equal to the sum of intensive and extensive growth, as shown below:

$$\begin{aligned} \text{INTESNIVE + EXTENSIVE} &= DN_{t0,2010} - DN_{t0,1996} + DN_{t1,2010} - DN_{t0,2010} \\ &= DN_{t1,2010} - DN_{t0,1996} \\ \text{INTESNIVE + EXTENSIVE} &= \text{NTLCHANGE} \end{aligned}$$

It may be argued, however, that both the total change in NTL intensity and the measure of extensive growth are overstated because they fail to account for the fact that there might have already existed some light in the newly added periphery. The following "corrected" measures of total change and extensive growth take account of such pre-existing lights:

$$\begin{aligned} \textbf{NTLCHGCORR} &= (DN_{t1,2010} - DN_{t0,1996}) - (DN_{t1,1996} - DN_{t0,1996}) \\ \textbf{EXTENCORR} &= (DN_{t1,2010} - DN_{t0,2010}) - (DN_{t1,1996} - DN_{t0,1996}) \end{aligned}$$

where, in each equation, the second expression in parenthesis on the right-hand side of the equation denotes the "correction factor", which is equal to the intensity of NTL that existed in the newly added periphery in 1996. Note that:

$$\text{NTLCHGCORR} = \text{INTENSIVE + EXTENCORR}$$

---

3. Methodology for Demarcating Urban Extents

There are several methodologies that have been devised over the years for utilizing NTL data for demarcating urban extents (see, for example, Imhoff *et al.*, 1997; Small *et al.*, 2011). For this product, we have developed a relatively straightforward method which utilizes a base land cover or built-up area layer for identification of pixel values that correspond to urban areas, and calculates an "optimal" country-specific NTL intensity (i.e. DN) threshold value to demarcate urban extents. In essence, the method defines urban extents by "calibrating" the NTL data against a higher-resolution land cover or built-up area map. The country-specific threshold is first calculated using 2010 global radiance calibrated NTL data. This threshold is then used to demarcate urban extents in both 2010 and 1996 on the assumption that the "optimal" NTL threshold in 1996 is the same as that derived for 2010.[21]

The global NTL urban extents and growth product utilizes the European Space Agency's 2009 Globcover classification[22] as its default base layer for the urban threshold calculation. However, the method for calculating the urban NTL threshold and defining urban extents is also perfectly compatible with any other layer that contains the necessary land class or built-up area information. NASA's MODIS Land Cover Type I product, the European Commission's Joint

[21] The adequacy of this assumption will be subject to testing in future product development. In general, we suspect that the optimal NTL threshold has increased gradually over time for any given country. If this is the case, the estimated growth of a given urban extent derived from applying the same threshold in 1996 as in 2010 will be biased downwards to some extent.

[22] http://due.esrin.esa.int/globcover/.

Research Center's Global Human Settlement Layer (GHSL),[23] and DLR's Global Urban Footprint (GUF)[24] are all examples of other layers that contain the necessary information and which could be used as alternative input into the method.

Below is a brief overview of the steps involved in the development of the global NTL urban extents and growth product:

1. For each country, extract the 2009 Globcover data and calculate associated NTL intensity values for each grid cell in 2010.
2. Tabulate NTL intensity values for pixels belonging to the urban and non-urban Globcover classes, and run a histogram analysis on the data.
3. Apply the threshold from the histogram analysis to the NTL data to derive the urban extents for both 2010 and 1996.

Prior to the processing of the NTL data, it is important to note that inter-annual calibration was performed using co-efficients provided by NOAA.[25] The purpose of this inter-annual calibration is to improve the comparability of the NTL data between different years. This is necessary to mitigate the effects of sensor degradation and changes in the operational settings of the DMSP-OLS satellites which, in the absence of inter-annual calibration, undermine the comparability of the data between different years.

3.1 Extract Globcover Data

For each country, Globcover data were extracted and classified into "urban" and "non-urban" land cover classes. More specifically, "urban" is taken as corresponding to land use class 190 ("Artificial surfaces and associated areas [urban areas > 50 %]") in the Globcover product, whilst all other land cover classes were aggregated into a single "non-urban" class. Each pixel in the Globcover data was then converted to a point with the associated ("urban" or "non-urban") land cover class. For every point, the corresponding NTL intensity (DN) value was determined using the global radiance calibrated NTL data for 2010.[26] Once each Globcover point was assigned an NTL intensity value, the urban threshold was calculated using the histogram analysis.

3.2 Histogram Analysis

To determine the "optimal" NTL threshold, the NTL intensity values for "urban" and "non-urban" were tabulated into histograms with bin sizes of 0.5. The histograms were compared by calculating the overall classification accuracy at each bin; the bin with the highest overall average accuracy was then chosen as the NTL threshold. Urban classification accuracy in this context represents the percentage of urban NTL pixels that are correctly identified as "urban" using a given assumed NTL intensity threshold – i.e. the percentage of pixels in all NTL defined

[23] http://ghslsys.jrc.ec.europa.eu/.

[24] http://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-9628/16557_read-40454/.

[25] See http://ngdc.noaa.gov/eog/dmsp/radcal_readme.txt.

[26] This data was extracted from the file "F16_20100111–20101209_rad_v4", which is downloadable from NOAA's website (http://ngdc.noaa.gov/eog/dmsp/download_radcal.html). The data from this file was chosen as it provides the closest temporal match to the Globcover product, which was produced in 2009.

urban extents whose classification matches that within the Globcover product.[27] Non-urban classification accuracy is similarly defined. Overall average accuracy of a given assumed NTL intensity threshold is the simple average of the urban and non-urban classification accuracies at that threshold.[28]

Table 2 illustrates the method using the example of Uganda. In this particular case, the "optimal" urban NTL intensity threshold which is selected for the delineating of urban extents corresponds to a DN value of 21. This is the NTL intensity threshold that yields the maximum average classification accuracy *vis-à-vis* the Globcover product. An important caveat to note is that the Globcover product (as with any remote sensing product) is itself subject to its own classification inaccuracies *vis-à-vis* actual land cover.[29] In this sense, the method for demarcating urban extents in the NTL data can only ever be as good as its default land cover or built-up area base layer. Improvements in this base layer will correspondingly result in improvements in the Global Night Time Lights Urban Extents and Growth Patterns product.

Table 2: Histogram analysis illustrating the selection of the optimal NTL urban threshold, example of Uganda

NTL intensity (DN)	Accuracy urban	Accuracy Non-urban	Average accuracy
18	96.43%	91.95%	94.19%
18.5	96.07%	92.44%	94.26%
19	95.78%	92.89%	94.34%
19.5	95.44%	93.30%	94.37%
20	95.15%	93.68%	94.41%
20.5	94.84%	94.02%	94.43%
21	94.53%	94.34%	94.43%
21.5	94.19%	94.63%	94.41%
22	93.83%	94.90%	94.36%
22.5	93.45%	95.15%	94.30%
23	93.07%	95.39%	94.23%
23.5	92.67%	95.60%	94.14%
24	92.23%	95.81%	94.02%
24.5	91.81%	96.00%	93.91%
25	91.36%	96.18%	93.77%
25.5	90.95%	96.35%	93.65%
26	90.48%	96.50%	93.49%

[27] This corresponds to the concept of "User's accuracy" in remote sensing, and provides information on errors of commission.

[28] For any given country, the number of non-urban pixels tends to be larger than the number of urban pixels. It follows that the method of selection of the "optimal" urban NTL intensity threshold implicitly places more weight on the accuracy of classification of urban pixels than it does the accuracy of classification of non-urban pixels. This is consistent with the fact that the product is one which is designed specifically with urban applications in mind.

[29] Globcover has an overall reported accuracy of 73 percent (see Defourny *et al*., 2009, for complete details of their accuracy assessment of Globcover).

NTL intensity (DN)	Accuracy urban	Accuracy Non-urban	Average accuracy
26.5	90.07%	96.65%	93.36%

3.3 Extract Urban Extents for 2010 and 1996

The urban threshold calculated from the histogram analysis was used to convert the input NTL data for 1996 and 2010 to binary raster data of above and below the threshold, thereby creating a binary "urban" / "non-urban" mask.[30] The resulting raster data were converted to polygons, and intersected with a default input layer of urban settlement / city points to associate names and population with the data. This default layer was taken to be CIESIN's GRUMP urban settlement layer, although, in principle, any user-defined input layer of settlement points could be used.[31] A intersect was then performed between the two years to determine how much each urban extent grew or shrunk, and whether an urban extent "appeared" or "disappeared" during the time period. In this context, an urban extent is classified as having "appeared" if it existed in 2010 but not in 1996, whilst it is classified as having "disappeared" if the opposite holds true. These transitions are captured by the variable "STATUS" in the summarized urban extents spreadsheet (see Table 1).

4. Limitations and Future Development

The version of the Global Night Time Lights Urban Extents and Growth Patterns product outlined in this technical documentation is an alpha version. As outlined in the introduction, the product is best suited to providing a broad overview of spatial patterns of urban growth and expansion within a country. Very much related to this, caution should be exercised in interpreting the precise quantitative estimates of, for example, the growth of individual urban extents that are reported in the Summarized Urban Extents spreadsheet. These estimates should be taken as providing more a general sense of the relative pace of expansion of different urban extents within a country. One use of the product in this sense is as a "first look" at patterns of urban expansion and growth within a country. Where interesting patterns are revealed, these can then be followed-up with more detailed analysis based on higher-resolution sources of satellite imagery (for example, medium-resolution Landsat imagery)[32] to provide more precise estimates of, for example, the pace of urban expansion. More generally, the product is likely to prove of most value when used in combination with other sources of spatial and urban-related data. The combination of data sources in analysis assists both in the triangulation of results and the

[30] The 1996 input NTL data was extracted from the file "F12_19960316–19970212_rad_v4" which is downloadable from NOAA's website (http://ngdc.noaa.gov/eog/dmsp/download_radcal.html).

[31] For example, in the SAR Urbanization Flagship work, a layer of all cities that had a *circa* 2000 population in excess of 100,000 was created from a variety of data sources (see CIESIN, 2013, for more details). In creating the layer of cities, it is important to have accurate information on longitude and latitude for each settlement point. This is necessary to ensure the accurate association of settlement points with urban extents.

[32] The archive of available Landsat imagery dates back to 1972 (For download options for Landsat Level 1 data products see http://landsat.usgs.gov//Landsat_Search_and_Download.php).

building-up of a more detailed picture of patterns of spatial and urban development within a country.

Future development of the product will focus on exploring and improving several features of the methodology outlined in this document. These include exploring the use of different base land cover or built-up area layers as input into the methodology for demarcating urban extents. Thus, although this product relies on ESA's Globcover 2009 product for its base layer, as mentioned above, the methodology is also perfectly compatible with other potential input layers which may hold potential advantages in terms of improving overall accuracy. Experimentation with, in particular, the Global Human Settlement Layer (GHSL), which is a built-up area layer that is attracting increasing attention from Bank teams, is planned in this regard.

Another important potential area of improvement that will be explored relates to the so-called "overglow" or "blooming" problem. Although there is not yet consensus within the remote sensing community on the exact relative importance of different potential causes of this problem,[33] concerns have been expressed by some researchers that the problem can confound, for example, quantitative estimates of extensive and intensive growth. The use of a calibrated threshold to demarcate urban extents helps to partially address some of the problems related to the overglow phenomenon. However, in an important recent development, Abrahams *et al*. (2014) have proposed a more rigorous "correction" to the problem based on the premise that overglow has its origins in distortions arising from the collection and processing of the data on-board the DMSP-OLS satellites. One avenue of future development, therefore, will be to explore whether this correction can be incorporated into the global product.

Finally, and most importantly, the future development of this product will be driven by feedback from Bank teams on its utility as a tool in policy dialogue work.

[33] Leading remote sensing experts tend to place relatively more weight on the hypothesis that overglow is a phenomenon driven, at least in part, by the atmospheric scattering of light from its point of origin. It is this atmospheric scattering, for example, which makes it difficult to view stars clearly at night except in relatively remote (non-built-up locations). Associated with this has been the use of NTL data in studies of the light pollution caused by urban areas. This is in addition to other sources which include, for example, geo-location uncertainty and sensor effects which are inherent in all optical systems.

References

Abrahams, A., N. Lozano-Gracia and C. Oram (2014), "Correcting Overglow in Nigthtime Lights Data", Unpublished manuscript.

Croft, T. (1978), "Night-time images of the Earth from Space", *Scientific American*, 239, pp 68 - 79.

Defourny, P. et al. (2009), "Accuracy assessment of a 300-m Global Land Cover Map: the GlobCover experience", Paper presented at the 33rd International Symposium on Remote Sensing of Environment (ISRSE), Stresa, Italy.

Doll, C.N.H. (2008), "CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications", *Center for International Earth Science Information Network (CIESIN)*, Columbia University, Palisades, New York.

Ellis, P.D. and M. Roberts (2015), *Leveraging Urbanization in South Asia: Managing Urban Development and Transformation for Improved Prosperity and Livability*, The World Bank, Washington, D.C.

Henderson, J.V., A. Storeygard and D. Weil (2011), "A Bright Idea for Measuring Economic Growth", *American Economic Review*, 101(3), pp 194-199.

Henderson, J.V., A. Storeygard and D. Weil (2012), "Measuring Economic Growth from Outer Space", *American Economic Review*, 102, pp 994-1028.

Imhoff, M.L., W.T. Lawrence, D.C. Stutzer, and C.D. Elvidge (1997), "A Technique for using Composite DMSP/OLS 'City Lights' Satellite Data to Map Urban Area", *Remote Sensing of Environment*, 61, pp 361–70.

Min, B. and K.M. Gaba (2014), "Tracking Electrification in Vietnam Using Nighttime Lights", *Remote Sensing*, 6(10), pp 9511-9529.

Nordhaus, W. and X. Chen (2014), "A Sharper Image? Estimates of the Precision of Nighttime Lights as a Proxy for Economic Statistics", *Journal of Economic Geography*, pp 1-30

Small, C., C. Elvidge, D. Balk and M. Montgomery (2011), "Spatial Scaling of Stable Night Lights", *Remote Sensing of Environment*, 115, pp 269-280.

Zhang, Q. and K. Seto (2011), "Mapping Urbanization Dynamics at Regional and Global Scales Using Multi-temporal DMSP/OLS Nighttime Light Data", *Remote Sensing of Environment*, 115, pp 2320-2329.

Zhou, N., K. Hubacek and M. Roberts (2015), "Analysis of spatial patterns of urban growth across South Asia using DMSP-OLS nighttime lights data", *Applied Geography*, pp 292-303.